Algebraic Algorithms for Cell Complexes

R. Laubenbacher
L. Garcia

Virginia Tech

B. Sturmfels
R. Datta

UC Berkeley

CARGO Incubation Grant

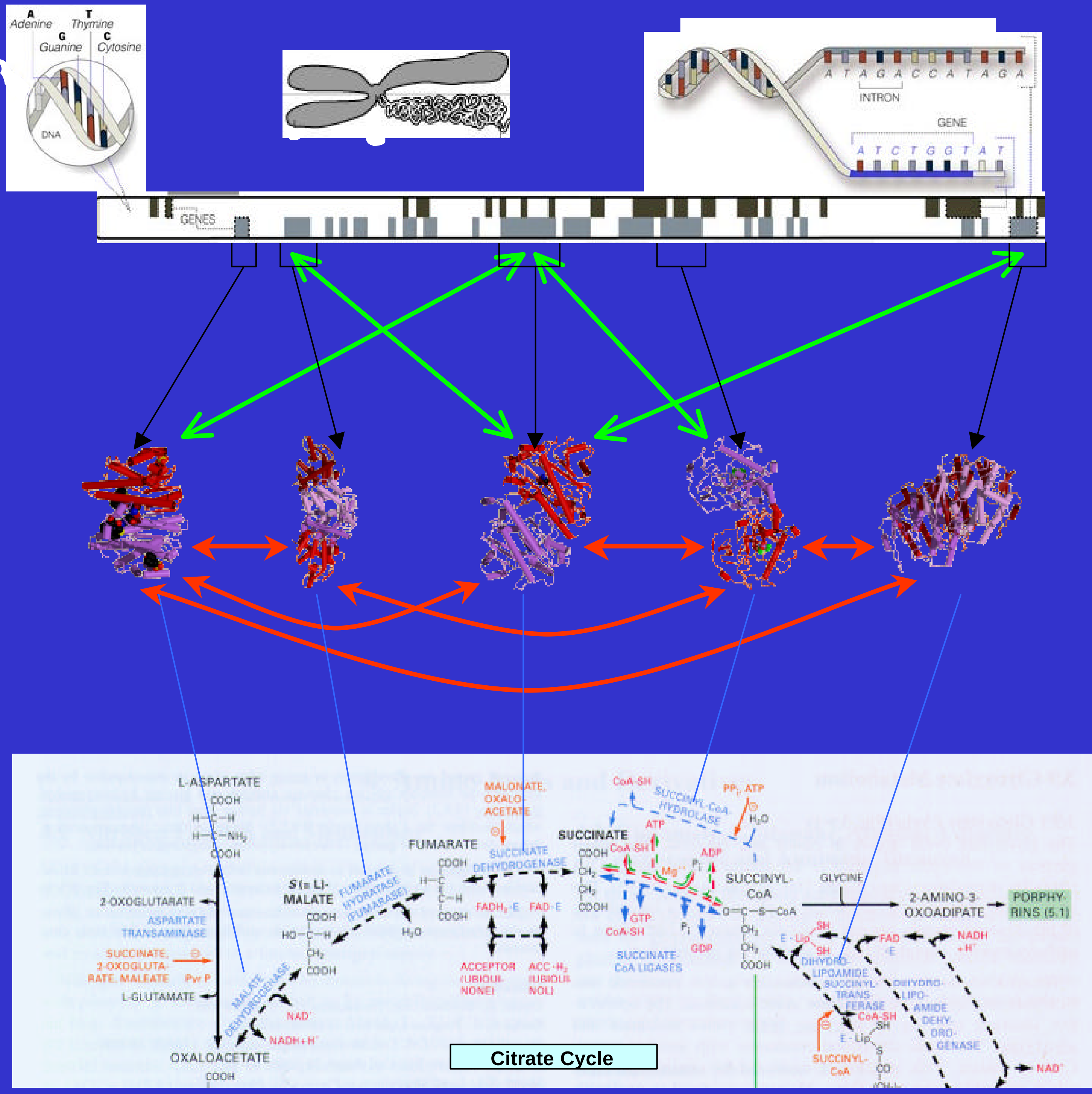
Objective: Develop practical algorithms for measures on large, **high dimensional** simplicial and cell complexes.

Motivation: Complexes occur as models in a variety of contexts. Focus on:

- Network dynamics.
- Free resolutions of monomial ideals in polynomial rings.

I. Interaction Patterns in Networks

- Gene regulatory networks
- Communications networks
- Social networks



GENOME

protein-gene interactions

PROTEOME

protein-protein interactions

METABOLISM

Bio-chemical reactions

Models

- Differential equations based
- Stochastic
- Boolean networks
- Various hybrids

New discrete model:

sequential dynamical systems

Sequential Dynamical Systems

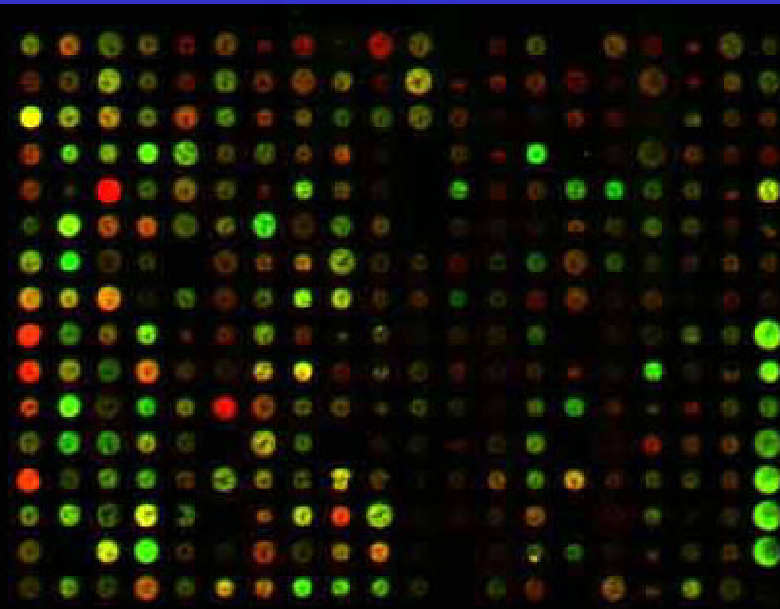
- Finite collection of nodes
- States taken from a finite field
- Graph connecting the nodes
- Local update functions of the nodes
- Update schedule

Compose local update functions according to update schedule to get discrete dynamical system

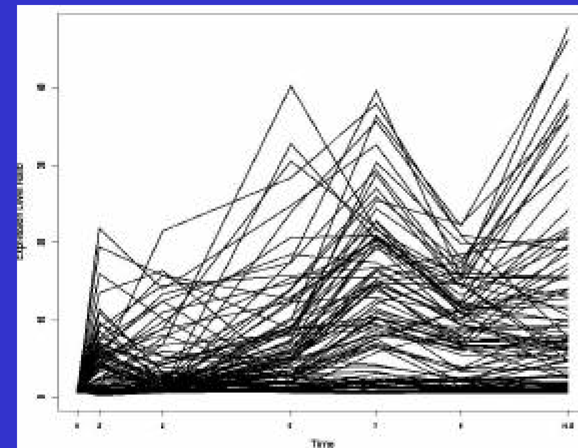
Features

- Update schedule allows different time scales for different nodes.
- Local update functions are polynomials.
- Larger state set allows better match of continuous data.
- Can introduce difference equations and thresholds.
- Can introduce stochastic features.

Modeling Approach



Gene Expression Microarray Image

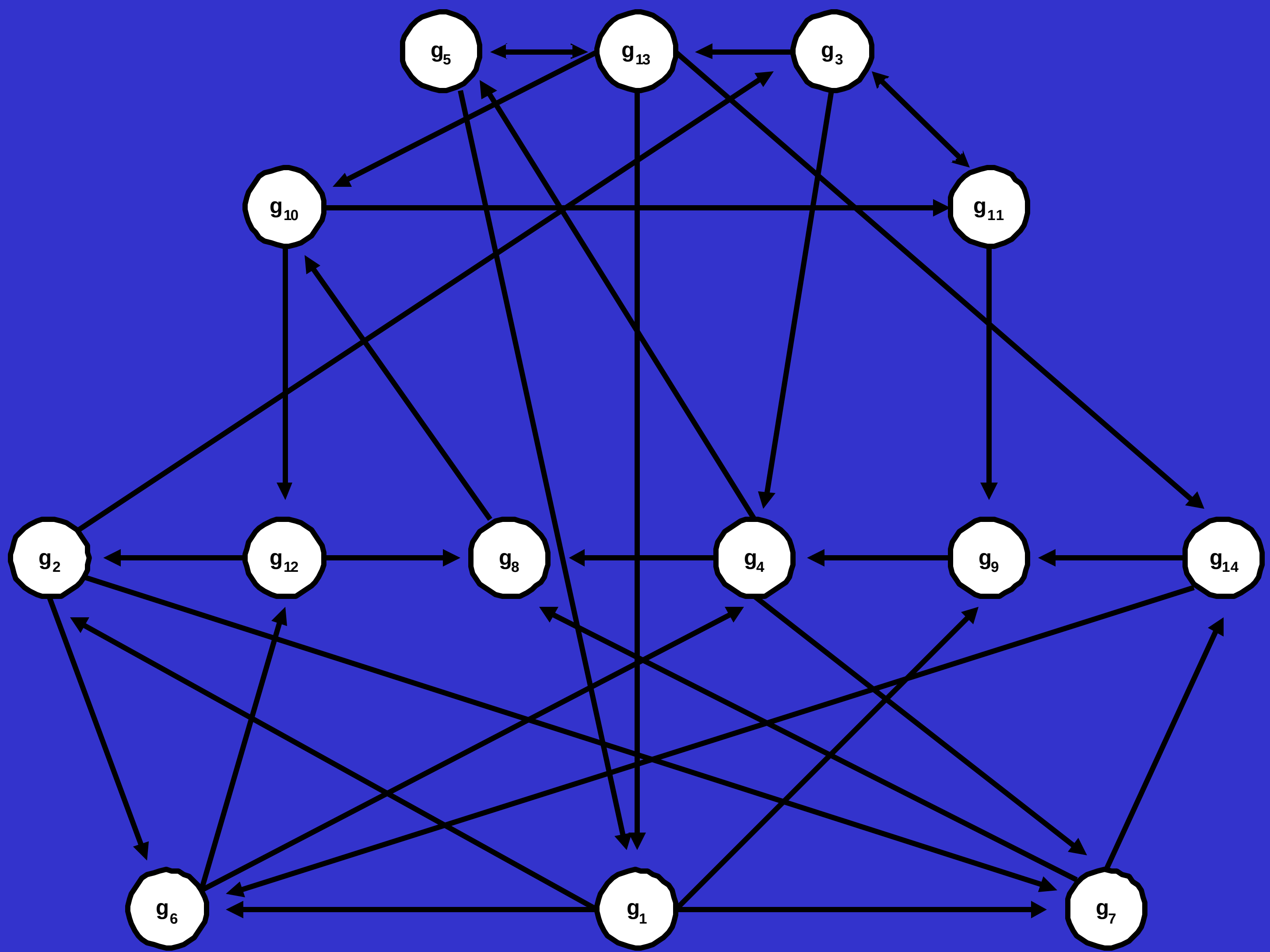


Gene Expression Profiles

Time Series of
Discrete States

Gene	GeneName	Time 0h	Time .5h	Time 2h	Time 5h	Time 7h	Time 9h	Time11.5
1	YBR166C	1	0	0	0	0	0	0
2	YGR066C	0	1	0	0	0	1	1
3	YMR048W	0	0	1	0	1	0	1
4	YGR035C	0	0	1	1	0	1	0
5	YHR172W	0	0	0	1	1	0	1
6	YDL085w	0	1	1	0	0	0	1
7	YCR005c	0	1	1	1	1	0	1
8	YBR001C	0	1	1	1	1	1	1
9	YER055c	0	1	1	0	1	0	0
10	YBL043W	0	1	1	1	1	1	0
11	YOR317W	0	1	0	1	0	0	0
12	YOL109W	0	0	0	0	1	1	1
13	YBL026W	0	0	0	0	0	1	0
14	YIL160C	0	1	0	0	0	0	1

One Possible Model of the Gene Regulatory Network



Search parameter space of SDS models for a given (small) network (with suitable restrictions) that match a given time series of expression data.

- Classify models based on common features:
  - similar promotion/suppression patterns (influence patterns)
  - similar local update functions
  - similar dynamics
  - similar network structure

Problem: need to define and measure SIMILAR

- Represent influence pattern as a partially ordered set (influence poset).
- Associate simplicial complexes to this poset:

  order complex: influence pathways

  covering complex: influence of genes

- Use topological/combinatorial measures on these complexes to define ‘similar influence pattern’

2. Communications Networks

Goal: Find measures for abnormal network behavior.

Network traffic: time series of directed weighted graphs.

Associate ‘neighborhood complex’ to a directed graph to get a time series of simplicial complexes.

Find suitable measures on the complexes.

Dimension of the complex as a distance measure.

Captures the same abnormal events as measure derived from the averaged sum of large eigenvalues of the Jacobian of the graph.

Also similar to edit distance.

3. Social Networks

Can socialization behavior of preschool children help predict future academic success?

Approach: Develop measures on interaction patterns.

The SMART Project

Date	Time	Period	ChildID	611	628	644	649	679	681	683	685	686	687	688	689	690	692	693	6
91700	9.21	1	611	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	
91700	11.08	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
91700	9.27	1	628	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
91700	11.17	1	628	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	
91700	11.26	1	644	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
91700	9.45	1	679	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
91700	9.43	1	681	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	
92700	9.04	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	9.16	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	9.46	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	10.2	1	611	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
92700	11.09	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	14.5	2	611	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
92700	14.52	2	611	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
92700	16.02	2	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	9.07	1	628	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	9.19	1	628	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	9.48	1	628	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	
92700	10.23	1	628	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92700	10.45	1	628	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
92800	10.41	1	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92800	14.47	2	611	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
92800	15.2	2	611	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
92800	10.45	1	628	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	
92800	9.57	1	644	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
92800	10.51	1	644	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

For each child construct a matrix:

Column labels: days in time period

Row labels: the other children

Construct a simplicial complex by interpreting the rows of this matrix as simplices.

Preliminary results: e.g., the dimension of the complex correlates extremely well with certain social and personal characteristics of the children.

4. Other Networks

- Message passing in CPU clusters.
- Infection patterns in populations
  (SDS-based LANL simulation)

Both lead to interaction posets, hence to simplicial complexes

II. Free Resolutions of Monomial Ideals.

- Use cell complexes to encode free resolutions of monomial ideals.
- Topological properties of the complex correspond to algebraic properties of the ideal.

Measures on Complexes

I. Homology

e.g., package of Dumas et.al. (integral coefficients), MAGMA, GAP, Groebner basis approach for coefficients in a field (e.g. CoCoa).

Compare, apply, and adapt software to our situation.

Explore other approaches (e.g., algebraic shifting).

II. A-homology

new homology (and homotopy) theory that is sensitive to the combinatorial connectivity of the complex, but is also robust.

An algorithm has been implemented for dimension one. Has proven useful, e.g., in measuring emergent properties in agent networks. Is also useful in studying the topology of monomial ideals and their complexes. General algorithm is possible, but needs to be implemented.

III. The f-vector

Integer vector that counts the number of simplices in each dimension.

Very important combinatorial invariant of simplicial complexes. No good algorithm exists at present.

Other Topics

- Computation of Nash equilibria.
- Log-linear models for higher-dimensional contingency tables.
- Cellular methods for solving algebraic equations.